Docket No.: 1509-253

PATENT #5

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of

Stephen John HINDE

Serial No. Not yet assigned : Group Art Unit: Not yet assigned

Filed: herewith : Examiner: N/A

For: SOUND EXCHANGES WITH VOICE SERVICE SYSTEMS

CLAIM OF PRIORITY AND
TRANSMITTAL OF CERTIFIED PRIORITY DOCUMENT

Assistant Commissioner For Patents
Washington, D.C. 20231

Dear Sir:

In accordance with the provisions of 35 U.S.C. 119, Applicant hereby claims the priority of:

Great Britain Patent Application No. 0030079.8 filed December 9, 2000

cited in the Declaration of the present application.

The certified copy is submitted herewith.

Respectfully submitted,

LOWE HAUPTMAN GILMAN & BERNER, LLP

Allan M. Lowe
Registration No. 19,641

LOWE HAUPTMAN GILMAN & BERNER, LLP *(22429)*
1700 Diagonal Road, Suite 310
Alexandria, Virginia 22314
(703) 684-1111 AML:klb
December 7, 2001

THIS PAGE BLANK (USPTO)

The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ

INVESTOR IN PEOPLE

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.
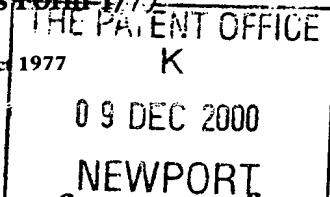
Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.
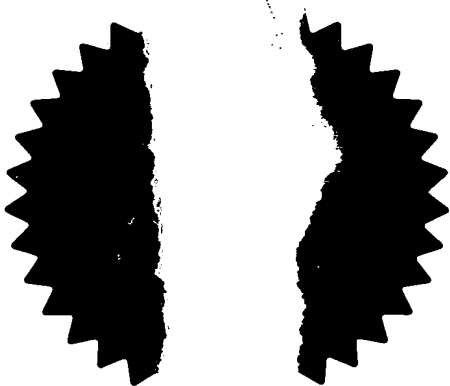
Signed

Dated   17 October 2001

THIS PAGE BLANK (USPTO)

THE PATENT OFFICE

K

0 9 DEC 2000

NEWPORT

The Patent Office

Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

The Patent Office

Cardiff Road
Newport
South Wales
NP10 8QQ

1. Your reference

30010815-GB

2. Patent application **0030079.8**
   (The Pa

9 DEC 2000

3. Full name, address and postcode of the or of each applicant *(underline all surnames)*

Hewlett-Packard Company
3000 Hanover Street
Palo Alto
CA 94304, USA

Patents ADP number *(if you know it)*

If the applicant is a corporate body, give the country/state of its incorporation

Delaware, USA 7812-985661

496588001

4. Title of the invention   Voice exchanges with voice service systems

5. Name of your agent *(if you have one)*

"Address for service" in the United Kingdom to which all correspondence should be sent *(including the postcode)*

Robert F Squibbs
Hewlett-Packard Ltd, IP Section
Filton Road
Stoke Gifford
Bristol BS34 8QZ

792818700\

Patents ADP number *(if you know it)*

6. If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and *(if you know it)* the or each application number	Country	Priority application number *(if you know it)*	Date of filing *(day / month / year)*

7. If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application	Number of earlier application		Date of filing *(day / month / year)*

8. Is a statement of inventorship and of right to grant of a patent required in support of this request? *(Answer 'Yes' if:*

   a) *any applicant named in part 3 is not an inventor, or*
   b) *there is an inventor who is not named as an applicant, or*
   c) *any named applicant is a corporate body.*
   See note (d))

   Yes

Patents Form 1/77

Patents Form 1/77

9. Enter the number of sheets for any of the
   following items you are filing with this form.
   Do not count copies of the same document

   Continuation sheets of this form

   Description   8

   Claim(s)   3

   Abstract   1

   Drawing(s)   3 + 3 e N

10. If you are also filing any of the following,
    state how many against each item.

    Priority documents

    Translations of priority documents

    Statement of inventorship and right
    to grant of a patent *(Patents Form 7/77)*

    Request for preliminary examination
    and search *(Patents Form 9/77)*   1

    Request for substantive examination
    (Patents Form 10/77)

    Any other documents   **fee sheet**
    (please specify)

11.   I/We request the grant of a patent on the basis of this application.

    Signature   T7. F. Smith,
    Robert Francis Squibbs   8 December 2000   Date

12. Name and daytime telephone number of
    person to contact in the United Kingdom   K Nommeots-Nomm   Tel: 0117-312-9947

Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

Notes

a)   *If you need help to fill in this form or you have any questions, please contact the Patent Office on 08459 500505.*

b)   *Write your answers in capital letters using black ink or you may type them.*

c)   *If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.*

d)   *If you have answered 'Yes' Patents Form 7/77 will need to be filed.*

e)   *Once you have filled in the form you must remember to sign and date it.*

f)   *For details of the fee and ways to pay please contact the Patent Office.*

Sound Exchanges With Voice Service Systems

Field of the Invention

The present invention relates to sound exchanges with voice service systems.

5

Background of the Invention

In recent years there has been an explosion in the number of services available over the World Wide Web on the public internet (generally referred to as the "web"), the web being composed of a myriad of pages linked together by hyperlinks and delivered by servers on request using the HTTP protocol. Each page comprises content marked up with tags to enable the receiving application (typically a GUI browser) to render the page content in the manner intended by the page author; the markup language used for standard web pages is HTML (HyperText Markup Language).

However, today far more people have access to a telephone than have access to a computer with an Internet connection. Sales of cellphones are outstripping PC sales so that many people have already or soon will have a phone within reach where ever they go. As a result, there is increasing interest in being able to access web-based services from phones. 'Voice Browsers' offer the promise of allowing everyone to access web-based services from any phone, making it practical to access the Web any time and any where, whether at home, on the move, or at work.

Human-to-human sound interaction is, in fact, far richer than the simple dialogs currently possible through the use of scripted voice service systems.

It is an object of the present invention to provide a method and system which enhances the available forms of sound interaction with a voice service system.

Summary of the Invention

According to one aspect of the present invention, there is provided a method of interacting with a human user through a sound service system, wherein the service system participates in a multi-turn sound exchange with the user, this sound exchange involving one or more

cycles in each of which the service and user take turns to provide a noise or utterance the form or content of which is already public.

Preferably, the service system also participates in normal voice dialog exchanges with the
5   human user, the service system using a respective manager for the normal voice dialogs and the multi-turn sound exchanges with control passing between the two managers as required, each manager when in control effecting this control according to a corresponding script.

10 The multi-turn sound exchange may include, or be constituted by, non-word sounds such as claps and whistles.

According to another aspect of the present invention, there is provided a sound service system comprising a sound input channel for receiving and interpreting sound input
15 signals, a sound output channel for generating sound output signals, and a multi-turn dialog manager connected to an output of the sound input channel and an input of the sound output channel, the multi-turn dialog manager being operative to manage the participation of the service system in a multi-turn sound exchange with a human user that involves one or more cycles in each of which the service and user take turns to provide a noise or
20 utterance the form or content of which is already public.

According to a further aspect of the present invention, there is provided a sound service system comprising:
    -   a sound input channel for receiving and interpreting sound input signals;
25 -   a sound output channel for generating sound output signals,
    -   a voice service manager connected to the output side of the sound input channel and the input side of the sound output channel, the voice service manager serving to manage normal voice dialog interactions with a human user;
    -   a multi-turn dialog manager connected to the output side of the sound input channel
30   and the input side of the sound output channel, the multi-turn dialog manager being operative to manage the participation of the service system in a multi-turn sound

exchange with a human user that involves one or more cycles in each of which the service and user take turns to provide a noise or utterance; and

- means for switching control between the voice service manager and the multi-turn dialog manager.

5

Further features of the invention are set out in the accompanying claims with the method features also being applicable to the system claims.

Brief Description of the Drawings

10 A method and system embodying the invention, for multi-turn sound exchanges with a sound service system, will now be described, by way of non-limiting example, with reference to the accompanying diagrammatic drawings, in which:

. **Figure 1**   is a diagram of a sound service system for effecting multi-turn dialog exchanges;

15 . **Figure 2**   is a diagram showing the sounds exchanged between a human user and the service system in respect of a first multi-turn sound exchange;

. **Figure 3**   is a diagram showing the sounds exchanged between a human user and the service system in respect of a second multi-turn sound exchange;

. **Figure 4**   is a diagram showing the sounds exchanged between a human user and the service system in respect of a third multi-turn sound exchange; and

20

. **Figure 5**   is a diagram showing a voice browser system including both a voice dialog manager and a multi-turn dialog manager.

Best Mode of Carrying Out the Invention

25 Figure 1 shows a sound service system for participating in a multi-turn sound exchange with a human user, this sound exchange involving one or more cycles in each of which the service and user take turns to provide a noise or utterance. Generally the form or content of the noise or utterance will already be public and known to the user so that they can participate fully in the exchange and gain a feeling of involvement.

30

The Figure 1 service system comprises a sound input channel 11 for receiving and interpreting sound inputs, a sound output channel 12 for generating sound output, and a

multi-turn dialog manager 10 connected to the output side of the sound input channel 11 and the input side of the sound output channel 12. The sound input channel comprises a microphone 13 feeding a speech recogniser 14 and a distinctive sound detection unit 15, this latter unit being designed to recognise specific non-word sounds such as handclaps and

5 whistles. The sound output channel 12 comprises a distinctive sound generator 16 for generating non-word sounds such as handclaps and whistles, a text-to-speech converter 17, an audio server 18 for outputting pre-recorded sound segments, and loudspeaker 19 for receiving the outputs of the generator 16, converter 17 and server 18 and generating corresponding sounds. The multi-turn dialog manager 10 is operative to manage the

10 participation of the service system in a multi-turn sound exchange, commanding the units of the output channel to generate appropriate sounds during the systems turn's in the multi-turn dialog and using the input channel to check for appropriate responses from the user. The multi-turn dialog manager 10 is, for example, arranged to interpret script files that define respective multi-turn dialogs.

15

Figure 2 illustrates a first multi-turn dialog, in this case initiated by the manager 10. The dialog proceeds as follows:

        Service: "<XXX product tune> + <whistle>"

        Human: "<whistles back>"

20   Service: "<Clapping sound>"

        Human: "<Claps>"

        Service: "XXX makes the best MP3 players ....<advertising>"

The distinctive sound generator 16 is used to generate the whistle and clap for the service system turns whilst the distinctive sound detection 15 unit detects these sounds repeated

25 back by the human user. The TTS converter 17 is used for generating the spoken words "XXX makes the best MP3 players". The audio server 18 is used to generate the XXX product tune and possibly also the advertising material (though this could be scripted for the TTS converter).

30 Figure 3 depicts a wholly speech-based multi-turn dialog running as follows:

        Service: "One , Two, Three.."

        Human: "Speak to me!"

Service: "And now its four.."

Human: "So tell me more!"

Service: <enters a normal dialog mode>

The multi-turn dialog actually terminates after the second human response. In the present

5   case, the service system now changes to a normal voice dialog mode under the control of a

voice dialog manager which can either be combined into one functional block with the

multi-turn dialog (MTD) manager 10, or embodied in a separate functional block (not

shown in Figure 1 but to be more fully described hereinafter with respect to the

embodiment of Figure 5).

10

The multi-turn dialogs illustrated in Figures 2 and 3 were both initiated by the MTD

manager 10 (for example, upon the user first contacting the service system). However, the

user may also initiate the dialog by uttering or otherwise generating the first sound

elements of the first turn of the dialog, this first turn being, now, a user turn. In this case,

15   the MTD manager 10 recognizes these first sound elements and participates in the

corresponding multi-turn dialog.

Figure 4 depicts a multi-turn dialog started by the user saying the word "wozsaar!" (in

other words: "What's up" meaning "What is happening?"). This is repeated back by the

20   service and this cycle the loops (repeats) either for as long as the user is willing to continue

or until a timeout period, or turn count, expires. In the illustrated case, the user ends the

multi-turn dialog by uttering an exit key word causing the service system to drop into a

normal voice dialog mode.

25   Figure 5 shows a voice browser 50 provided with multi-turn dialog capability. Voice

browser 50 is located in the communications infrastructure (being, for example, provided

by a PSTN or PLMN operator or by an ISP). A voice browser allows people to access the

Web using speech and is interposed between equipment 41 of a user 42, and a voice page

server 40. This server 40 holds voice service pages (text pages) that are marked-up with

30   tags of a voice-related markup language (or languages). When a normal voice dialog page

(such as page 26) is requested by the user 42, it is interpreted at a top level (dialog level) by

a voice dialog manager 22 of the voice browser 33 and output intended for the user is

passed in text form to the output channel 12 of the browser. The output channel is here shown as comprising a language generator 30 for driving a Text-To-Speech (TTS) converter 17 and a distinctive sound generator 16. The output of channel 12 is passed over a sound connection (such as a telephone voice circuit or VoIP connection) to the user

5 equipment 41. User voice (and other sound) input is passed back over this connection to an input channel 11 of the voice browser. The input channel in this case comprises a speech recognition unit 14 and a distinctive sound detection unit 15, both feeding a language understanding unit 21. The input channel 11 uses lexicon and grammar data 25 to determine what sounds/words have been received and may seek to understand this input in

10 the context of what has gone before. For normal voice dialog operation, the output of the channel 11 is passed to the voice dialog manager 22 which then determines what action is to be taken according to the received input and the directions in the original page.

The voice browser 50 further comprises a MTD manager 23 that is a distinct functional

15 element from the normal voice dialog manager 22, the MTD manager operating according to one or more multi-turn dialog scripts 23. The scripts are, for example, loaded into the MTD manager 23 when the voice browser first contacts a voice website hosted on server 40, these scripts being retained whilst the voice browser remains at the same voice website (in contrast, as the browser moves between pages of the website, different scripts 26 are

20 loaded into and out of the voice dialog manager 22).

The voice browser can operate in two modes, namely a normal voice dialog mode in which the voice dialog manager 22 is in control and runs its currently loaded script 26, and an MTD mode in which the MTD manager 23 is in control and runs a selected one of its

25 currently loaded scripts 27. The current mode of the browser is held by mode unit 24 which indicates the current mode of the voice browser to the managers 22, 23, the language understanding unit, and the language generator 30.

When the browser is in its normal voice dialog mode, the language understanding unit 21

30 uses the grammar appropriate to script 26 but, in addition, is caused to look out for user input sound elements that correspond to the initial elements of any of the multi-turn dialog script starting with a user turn. If the initial elements of such a MTD script are detected in

the input sound stream, the unit 21 changes the mode of the browser held by mode unit 24 to the MTD mode. As a result, MTD manager 23 assumes control of the following sound exchange which it controls in accordance with the script 27 whose initial elements were detected.

5

The MTD mode can also be entered by the current script 26 requesting the unit 24 to change modes, the script 26 also informing the manager 23 which multi-turn dialog script 27 is to be executed.

10 When the browser is in its MTD mode, the language understanding unit 21 uses the grammar appropriate to the selected script 27 but, in addition, is caused to look out for user input sound elements that correspond to exit key words or phrases indicating that the user wishes to terminate the current multi-turn dialog exchange. If an exit key word or phrase is detected in the input sound stream, the unit 21 changes the mode of the browser 15 held by mode unit 24 to the normal voice dialog mode. As a result, voice dialog manager 22 assumes control of the following sound exchange in accordance with script 26.

The normal voice dialog mode can also be entered as a result of the MTD manager causing the unit 24 to change the set mode upon the MTD manager finishing execution of a current 20 multi-turn dialog script 27.

During the course of execution of a multi-turn dialog script, the user's input can be treated in a number of different ways:

(a) - checked only for occurrence of a user input, not necessarily as expected;
25 (b) - checked for expected user input;
(c) - checked for one of several possible expected user inputs.

In respect of (b), if the expected user input is not received, the multi-turn dialog can be terminated or a correction dialog entered. In the case of (c), the identity of received input (if one of the expected inputs) can be used to determine which of several branches in the 30 dialog is pursued by the MTD manager 23; alternatively, the identity of the received input can be used to select a particular voice dialog script 26 to be used when the normal voice

dialog mode is entered at the end of the current multi-turn dialog, this identity being passed to manager 22.

Many variants are, of course, possible to the arrangements described above. For example, the MTD manager can be provided with functionality for ensuring that the service system executes its next turn promptly on the conclusion of the user's turn. This functionality can include means for predicting when the user's input will terminate having regard to its speed of delivery. Such promptness of response increases the user's feeling of interaction with the service system.

Whilst the turns of the multi-turn dialogs will generally be public knowledge, this is not always necessary to achieve a feeling of involvement.

It will be appreciated by persons skilled in the art that the managers 22 and 23 will normally be implements in software and can, in fact, simply be separate instantiations of the same manager class.

CLAIMS

1. A method of interacting with a human user through a sound service system, wherein the service system participates in a multi-turn sound exchange with the user, this sound exchange involving one or more cycles in each of which the service and user take turns to provide a noise or utterance the form or content of which is already public.

2. A method according to claim 1, wherein the multi-turn sound exchange is initiated by the service system.

3. A method according to claim 1, wherein the multi-turn sound exchange is initiated by the human user.

4. A method according to claim 1, wherein the service system also participates in normal voice dialog exchanges with the human user, the service system using the same dialog manager for the normal voice dialogs and the multi-turn sound exchanges with each being effected according to a corresponding script run by the dialog control as required.

5. A method according to claim 1, wherein the service system also participates in normal voice dialog exchanges with the human user, the service system using a respective manager for the normal voice dialogs and the multi-turn sound exchanges with control passing between the two managers as required, each manager when in control effecting this control according to a corresponding script.

6. A method according to claim 5, including the step of the user inputting a sound corresponding to the start of a particular multi-turn sound exchange whilst the voice dialog manager is in control, the service system recognising this sound and putting the multi-turn dialog manager in control to run the script corresponding to said particular multi-turn sound exchange.

7. A method according to claim 6, wherein the service system is adapted to recognise and distinguish between sounds corresponding to multiple different multi-turn sound

exchanges.

8. A method according to claim 5, including the step of the user inputting a sound, whilst the multi-turn dialog manager is in control, indicative that the user wishes to exit the current multi-turn sound exchange, the service system recognising this sound and putting the voice dialog manager in control to run an appropriate voice dialog script.

9. A method according to claim 5, wherein the scripts for the voice dialog manager and multi-turn dialog manager are independently loaded.

10. A method according to claim 5, wherein the voice service system comprises a voice browser for interpreting scripts provided by voice sites hosted by page servers, one or more multi-turn sound exchange scripts being loaded to the multi-turn dialog manager upon a user first contacting a said voice site and remaining loaded whilst the user browses the voice pages of the site, the currently-visited voice page of the site being loaded to the voice dialog manager.

11. A method according to any one of the preceding claims, wherein the multi-turn sound exchange includes, or is constituted by, non-word sounds.

12. A method according to any one of the preceding claims, wherein the multi-turn sound exchange is of a looping nature and terminates in response to at least one of:
- explicit user request;
- timeout of a predetermined time from commencement of the exchange;
- execution of a preset number of cycles.

13. A method according to any one of the preceding claims, wherein the user's input during at least one turn of the multi-turn sound exchange, is used to determine which of two or more branches in the service system's part of the multi-turn sound exchange is taken by the service system.

14. A method according to any one of the preceding claims, wherein the user's input

during at least one turn of the multi-turn sound exchange, is used to determine the identity of a voice dialog script followed by the service system following termination of the multi-turn sound exchange.

5   **15.** A sound service system comprising a sound input channel for receiving and interpreting sound input signals, a sound output channel for generating sound output signals, and a multi-turn dialog manager connected to an output of the sound input channel and an input of the sound output channel, the multi-turn dialog manager being operative to manage the participation of the service system in a multi-turn sound exchange with a

10   human user that involves one or more cycles in each of which the service and user take turns to provide a noise or utterance the form or content of which is already public.

   16. A sound service system comprising:

-   a sound input channel for receiving and interpreting sound input signals;

15   -   a sound output channel for generating sound output signals,

-   a voice service manager connected to the output side of the sound input channel and the input side of the sound output channel, the voice service manager serving to manage normal voice dialog interactions with a human user;

-   a multi-turn dialog manager connected to the output side of the sound input channel

20   and the input side of the sound output channel, the multi-turn dialog manager being operative to manage the participation of the service system in a multi-turn sound exchange with a human user that involves one or more cycles in each of which the service and user take turns to provide a noise or utterance; and

-   means for switching control between the voice service manager and the multi-turn

25   dialog manager.

   17. A system according to claim 15 or claim16, wherein the multi-turn sound exchange includes, or is constituted by, non-word sounds, the system including specific means for recognising and/or generating said non-word sounds.

30

ABSTRACT

5   **Sound Exchanges With Voice Service Systems**

A sound service system (50) participates in a multi-turn sound exchange with a human user (42), this sound exchange involving one or more cycles in each of which the service and user take turns to provide a noise or utterance the form or content of which is already

10   public. The service system (50) preferably also participates in normal voice dialog exchanges with the human user, the service system using a respective manager for the normal voice dialogs (22) and the multi-turn sound exchanges (23) with control passing between the two managers (22,23) as required, each manager when in control effecting this control according to a corresponding script (26,27).

15

(Fig. 5)
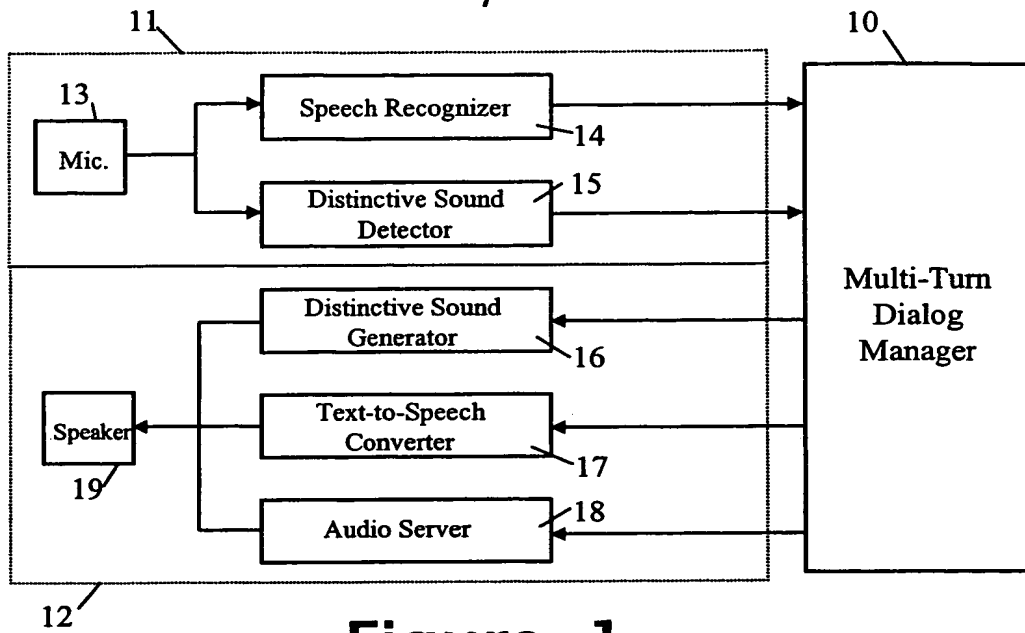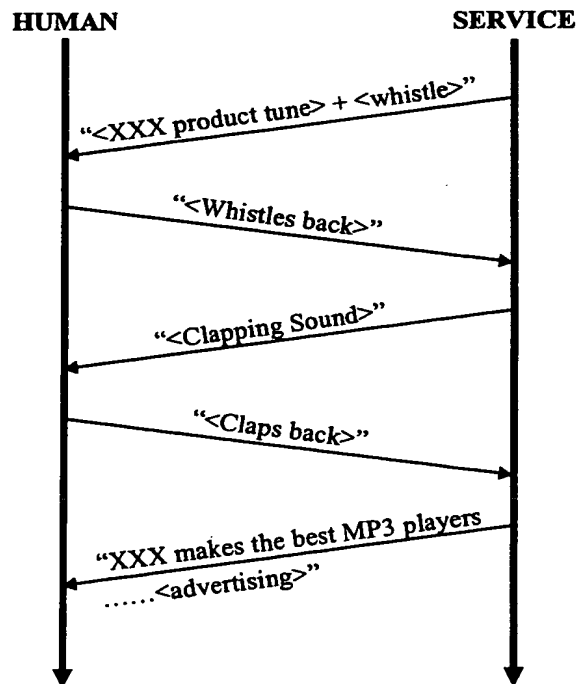
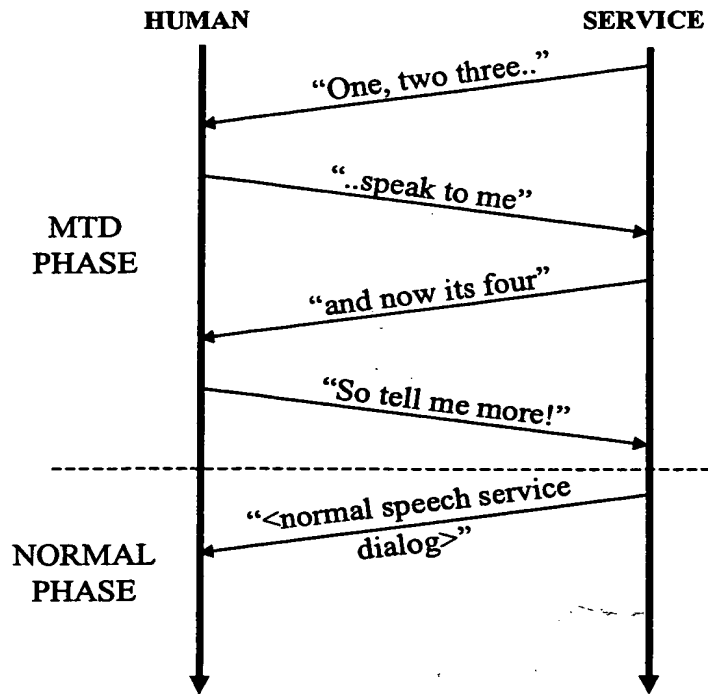**Figure 1**

Figure 2

THIS PAGE BLANK (USPTO)

HUMAN   SERVICE

"One, two three.."

"..speak to me"

MTD
PHASE

"and now its four"

"So tell me more!"

"<normal speech service dialog>"

NORMAL
PHASE

Figure 3

HUMAN   SERVICE

"Wozsaaar!"

"Wozsaaar!"

MTD
PHASE

"Wozsaaar!"

"Wozsaaar!"

"<exit command>"
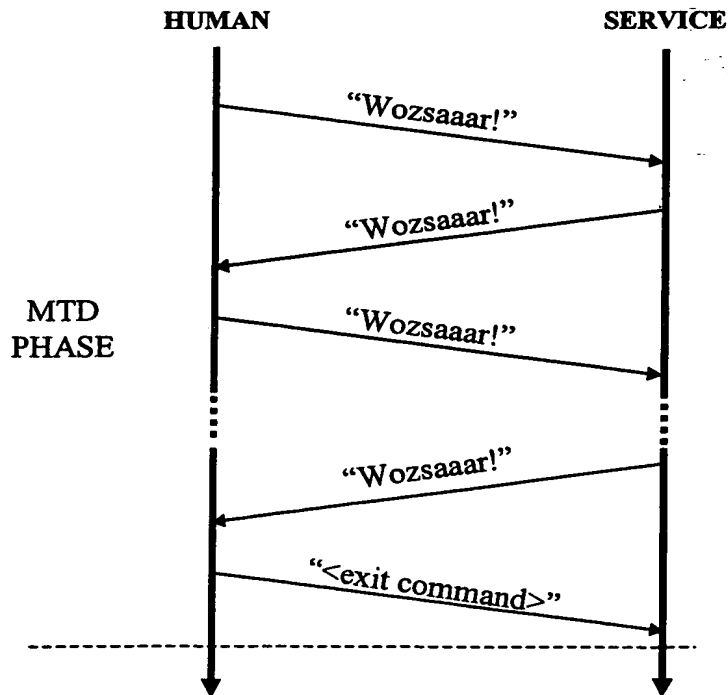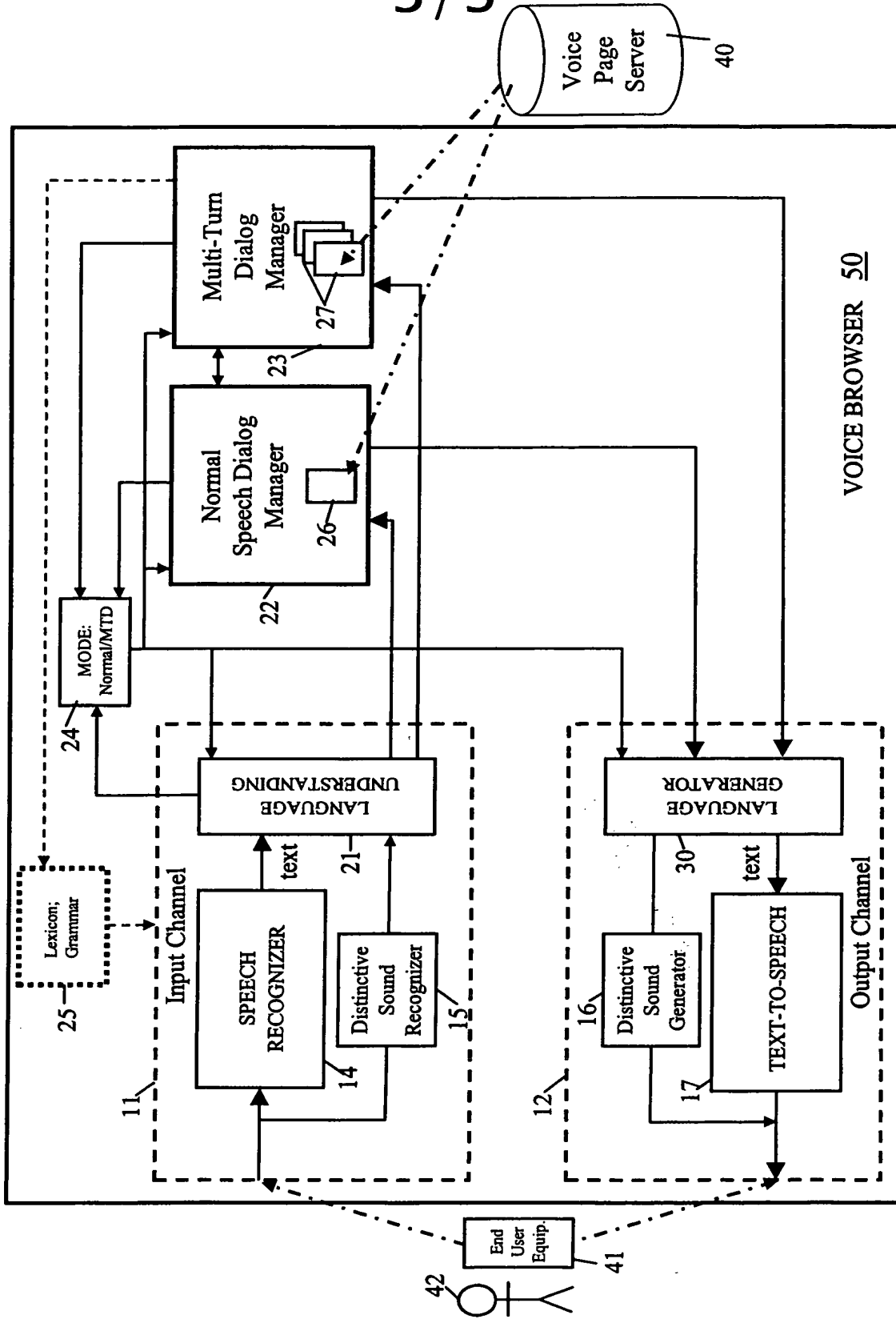
Figure 4

Figure 5

THIS PAGE BLANK (USPTO)

INVESTOR IN PEOPLE

Patents Act 1977
Search Report under Section 17

Databases searched:

UK Patent Office collections, including GB, EP, WO & US patent specifications, in:

    UK Cl (Ed.S): H4L (LEUF, LRAB), H4K (KOD4), G4F

    Int Cl (Ed.7): G06F 17/30, G08B 3/00, /02, /10, H04B 11/00, H04Q7/22, H04M 1/22, 3/493

Other: Online: WPI, EPODOC, JAPIO, INSPEC

Documents considered to be relevant:

Category	Identity of document and relevant passage	Relevant to claims
X	EP 0992980 A2 (LUCENT) See abstract	1-4, 11-17
X	EP 0878948 A2 (AT&T) See abstract	1-4, 11-17
X	WO 00/64137 A1 (NABHA) See page 3 lines 9-30	1-4, 11-17
X	WO 00/62278 A1 (INROAD) See page 3 lines 17-32	1-4, 11-17
X	WO 00/05708 A1 (MOTOROLA) See abstract	1-4, 11-17
X	WO 99/57714 A1 (GENERAL MAGIC) See abstract	1-4, 11-17